Multi-bosonic algorithms for dynamical fermion simulations[*]

I. Montvay

Deutsches Elektronen-Synchrotron DESY,
Notkestr. 85, D-22603 Hamburg, Germany
E-mail: istvan.montvay@desy.de

1 Introduction

The numerical simulation of quantum field theories with fermions is an interesting and difficult computational problem. The basic difficulty is the necessary calculation of very large determinants, the determinants of *fermion matrices*, which can only be achieved by some stochastic procedure with the help of auxiliary bosonic "pseudofermion" fields.

In hybrid Monte Carlo algorithms the number of pseudofermion fields corresponds to the number of fermion field components. The evolution of the pseudofermion fields in the updating process is realized by discretized molecular dynamics equations. The error implied by the finite length of discretization steps is corrected for by a global accept-reject decision which involves a fermion matrix inversion. The ingredients of the *two-step multi-bosonic algorithm* [1], which will be considered in this contribution, are somewhat different but still in a general sense similar. The number of pseudofermion fields is multiplied by the order of a polynomial approximation of some negative power $x^{-\alpha}$ of the fermion matrix. These auxiliary bosonic fields are updated acording to a *multi-bosonic action* [2] by using simple methods known from bosonic quantum field theories, as heatbath and

---

[*]Talk given at the workshop on *Molecular Dynamics on Parallel Computers*, February 1999, NIC, Jülich; to appear in the proceedings.

overrelaxation. The error of the polynomial approximation is corrected also here in a global accept-reject decision. This is realized in a so called *noisy correction step* [3] by using a better polynomial approximation, which realizes a kind of generalized inversion of the fermion matrix.

In my talk I review some recent developments of the multi-bosonic algorithms. The performance of the two-step multi-bosonic algorithm is illustrated in a recent large scale Monte Carlo simulation of the supersymmetric Yang-Mills theory [4, 5], which is being performed at NIC, Jülich. This application shows that the algorithm is able to cope with the difficulties arising at nearly zero gluino masse in reasonably large physical volumes.

2 Multi-bosonic actions

The multi-bosonic representation of the fermion determinant [2] is based on the approximation

$$|\det(Q)|^{N_f} = \left\{\det(Q^\dagger Q)\right\}^{N_f/2} \simeq \frac{1}{\det P_n(Q^\dagger Q)} \ , \tag{1}$$

where N_f (Dirac-) fermion flavours are considered and the polynomial P_n satisfies

$$\lim_{n\to\infty} P_n(x) = x^{-N_f/2} \tag{2}$$

in an interval $x \in [\epsilon, \lambda]$. This interval is chosen such that it covers the spectrum of the squared hermitian fermion matrix $\tilde{Q}^2$. The hermitian matrix $\tilde{Q}$ is defined from the original fermion matrix Q

$$\tilde{Q} \equiv \gamma_5 Q = \tilde{Q}^\dagger \tag{3}$$

which is assumed here to satisfy the relation

$$Q^\dagger = \gamma_5 Q \gamma_5 \ . \tag{4}$$

Note that in (1) only the absolute value of the determinant is taken which leaves out its sign (or phase). This can be taken into account at the evaluation of expectation values by reweighting.

For the multi-bosonic representation of the determinant one uses the roots of the polynomial r_j, $(j = 1, \ldots, n)$

$$P_n(Q^\dagger Q) = P_n(\tilde{Q}^2) = r_0 \prod_{j=1}^{n} (\tilde{Q}^2 - r_j) \ . \tag{5}$$

Assuming that the roots occur in complex conjugate pairs, one can introduce the equivalent forms

$$P_n(\tilde{Q}^2) = r_0 \prod_{j=1}^{n} [(\tilde{Q} \pm \mu_j)^2 + \nu_j^2] = r_0 \prod_{j=1}^{n} (\tilde{Q} - \rho_j^*)(\tilde{Q} - \rho_j) \tag{6}$$

where $r_j \equiv (\mu_j + i\nu_j)^2$ and $\rho_j \equiv \mu_j + i\nu_j$. With the help of complex scalar (pseudofermion) fields Φ_{jx} one can write

$$\det P_n(Q^\dagger Q)^{-1} \propto \prod_{j=1}^{n} \det[(\tilde{Q} - \rho_j^*)(\tilde{Q} - \rho_j)]^{-1}$$

$$\propto \int [d\Phi] \, \exp\left\{ -\sum_{j=1}^{n} \sum_{xy} \Phi_{jy}^+ \, [(\tilde{Q} - \rho_j^*)(\tilde{Q} - \rho_j)]_{yx} \, \Phi_{jx} \right\} . \tag{7}$$

The exponent in (7) is the (negative) multi-bosonic action. Since it is quadratic in $\tilde{Q}$, its locality properties are inherited from the fermion matrix Q. For instance, if Q has only nearest neighbour interactions then the multi-bosonic action (7) extends up to next-to-nearest neighbours.

The multi-bosonic representation of the fermion determinant (7) can be used for a Monte Carlo procedure in terms of the pseudofermion fields Φ_{jx}. The difficulty for small fermion masses in large physical volumes is that the *condition number* λ/ϵ becomes very large ($10^4 - 10^6$) and very high orders $n = \mathcal{O}(10^3)$ are needed for a good approximation. This requires large storage and the autocorrelation becomes bad since it is proportional to n. An additional question is how to control the systematic errors introduced by the polynomial approximation in (1). In principle one has to perform the limit to infinite order of the approximation $n \to \infty$, which means in practical terms a need to investigate the dependence of the results on n.

Several solutions for eliminating the systematical errors due to the finite order of approximation n in (1) are possible. For instance, one can calculate a correction factor from the eigenvalues of $\tilde{Q}^2$ and introduce a corresponding global Matropolis accept-reject step in the updating [6]. The necessary calculation of the eigenvalues leads, however, to an algorithm growing with the square of the number of lattice points.

A better solution is to apply a noisy correction step [3] which is especially simple in case of $N_f = 2$ flavours when an iterative inversion is sufficient [7]. This can be generalized to an arbitrary number of flavours in the two-step multi-bosonic scheme [1] which will be discussed in detail in the next section. The special case of $N_f = 1$ flavours can be dealt with in a non-hermitian version [7, 8] applied directly to Q, insted of $Q^\dagger Q$ in (1). This works well for heavy fermion masses [9] when the spectrum of Q can be covered by en ellipse but it would be very cumbersome for small fermion masses where the eigenvalues are surrounding zero.

Another possibility to perform the corrections of the systematic errors of the polynomial approximation is reweighting in the expectation values. For the special case of $N_f = 2$ this has been introduced in the polynomial hybrid Monte Carlo scheme [10]. The case of arbitrary N_f can be solved by an appropriate polynomial approximation, which can also be implemented in the two-step multi-bosonic approach (see section 3.2).

3

Other approaches for eliminating the systematic errors are possible, for instance, by choosing some specific ways of optimizing the approximate polynomials (see [11, 12]). It is also possible to combine the multi-bosonic idea with other methods of dynamical fermion simulations. To review all attempts would take too much time and most of the proposals were not yet tested in really large scale simulations. In what follows I shall concentrate on the two-step multi-bosonic scheme which proved to be efficient in recent simulations of SU(2) Yang-Mills theory with light gluinos [4, 5].

3   Two-step multi-bosonic algorithm

The dynamical fermion algorithm directly using the multi-bosonic representation in (7) has difficulties with large storage requirements and long autocorrelations. One can achieve substantial improvements on both these problems by introducing a two-step polynomial approximation [1, 13]. In this two-step approximation scheme (2) is replaced by

$$\lim_{n_2 \to \infty} P_{n_1}^{(1)}(x) P_{n_2}^{(2)}(x) = x^{-N_f/2} , \qquad x \in [\epsilon, \lambda] . \tag{8}$$

The multi-bosonic representation is only used for the first polynomial $P_{n_1}^{(1)}$ wich provides a first crude approximation and hence the order n_1 can remain relatively low. The correction factor $P_{n_2}^{(2)}$ is realized in a stochastic *noisy correction step* with a global accept-reject decision during the updating process (see section 3.1). In order to obtain an exact algorithm one has to consider in this case the limit $n_2 \to \infty$. For very small fermion masses it turned out more practicable to fix some large n_2 and perform another small correction in the evaluation of expectation values by *reweighting* with a still finer polynomial (see section 3.2).

3.1   Update correction: global accept-reject

The idea to use a stochastic correction step in the updating [3], instead of taking very large polynomial orders n, was proposed in the case of $N_f = 2$ flavours in [7]. $N_f = 2$ is special because the function to be approximated is just x^{-1} and $P_{n_2}^{(2)}(x)$ can be replaced by the calculation of the inverse of $x P_{n_1}^{(1)}(x)$. For general N_f one can take the two-step approximation scheme introduced in [1] where the two-step multi-bosonic algorithm is described in detail. The theory of the necessary optimized polynomials is summarized in section 4 following [13].

In the two-step approximation scheme for N_f flavours of fermions the absolute value of the determinant is represented as

$$|\det(Q)|^{N_f} \simeq \frac{1}{\det P_{n_1}^{(1)}(\tilde{Q}^2) \det P_{n_2}^{(2)}(\tilde{Q}^2)} . \tag{9}$$

4

The multi-bosonic updating with n_1 scalar pseudofermion fields is performed by heatbath and overrelaxation sweeps for the scalar fields and Metropolis sweeps for the gauge field. After a Metropolis sweep for the gauge field a global accept-reject step is introduced in order to reach the distribution of gauge field variables $[U]$ corresponding to the right hand side of (9). The idea of the noisy correction is to generate a random vector η according to the normalized Gaussian distribution

$$\frac{e^{-\eta^\dagger P_{n_2}^{(2)}(\tilde{Q}[U]^2)\eta}}{\int[d\eta]e^{-\eta^\dagger P_{n_2}^{(2)}(\tilde{Q}[U]^2)\eta}} \ , \tag{10}$$

and to accept the change $[U'] \leftarrow [U]$ with probability

$$\min\left\{1, A(\eta;[U'] \leftarrow [U])\right\} \ , \tag{11}$$

where

$$A(\eta;[U'] \leftarrow [U]) = \exp\left\{-\eta^\dagger P_{n_2}^{(2)}(\tilde{Q}[U']^2)\eta - \eta^\dagger P_{n_2}^{(2)}(\tilde{Q}[U]^2)\eta\right\} \ . \tag{12}$$

The Gaussian noise vector η can be obtained from η' distributed according to the simple Gaussian distribution

$$\frac{e^{-\eta'^\dagger \eta'}}{\int[d\eta']e^{-\eta'^\dagger \eta'}} \tag{13}$$

by setting it equal to

$$\eta = P_{n_2}^{(2)}(\tilde{Q}[U]^2)^{-\frac{1}{2}}\eta' \ . \tag{14}$$

In order to obtain the inverse square root on the right hand side of (14), we can proceed with polynomial approximations in two different ways. The first possibility was proposed in [1] with $x \equiv \tilde{Q}^2$ as

$$P_{n_2}^{(2)}(x)^{-\frac{1}{2}} \simeq R_{n_3}(x) \simeq x^{N_f/4}S_{n_s}[P_{n_1}^{(1)}(x)] \ . \tag{15}$$

Here

$$S_{n_s}(P) \simeq P^{\frac{1}{2}} \tag{16}$$

is an approximation of the function $P^{\frac{1}{2}}$ on the interval $P \in [\lambda^{-N_f/2}, \epsilon^{-N_f/2}]$. The polynomial approximations R_{n_3} and S_{n_s} can be determined by the same general procedure as $P_{n_1}^{(1)}$ and $P_{n_2}^{(2)}$. It turns out that these approximations are "easier" in the sense that for a given order higher precisions can be achieved than, say, for $P_{n_1}^{(1)}$.

Another possibility to obtain a suitable approximation for (14) is to use the second decomposition in (6) and define

$$P_{n_2}^{(1/2)}(\tilde{Q}) \equiv \sqrt{r_0}\prod_{j=1}^{n_2}(\tilde{Q} - \rho_j) \ , \qquad P_{n_2}^{(2)}(\tilde{Q}^2) = P_{n_2}^{(1/2)}(\tilde{Q})^\dagger P_{n_2}^{(1/2)}(\tilde{Q}) \ . \tag{17}$$

5

Using this form, the noise vector η necessary in the noisy correction step can be generated from the gaussian vector η' according to

$$\eta = P_{n_2}^{(1/2)}(\tilde{Q})^{-1}\eta' \; , \tag{18}$$

where $P_{n_2}^{(1/2)}(\tilde{Q})^{-1}$ can be obtained as

$$P_{n_2}^{(1/2)}(\tilde{Q})^{-1} = \frac{P_{n_2}^{(1/2)}(\tilde{Q})^{\dagger}}{P_{n_2}^{(2)}(\tilde{Q}^2)} \simeq P_{n_3}(\tilde{Q}^2)P_{n_2}^{(1/2)}(\tilde{Q})^{\dagger} \; . \tag{19}$$

In the last step P_{n_3} denotes a polynomial approximation for the inverse of $P_{n_2}^{(2)}$ on the interval $[\epsilon, \lambda]$. Note that this last approximation can also be replaced by an iterative inversion of $P_{n_2}^{(2)}(\tilde{Q}^2)$. However, tests showed that the inversion by a least-squares optimized polynomial approximation is much faster because, for a given precision, less matrix multiplications have to be performed.

In the simulation with light dynamical gluinos [4, 5] mainly the second form in (18)-(19) has been used. The first form could, however, be used as well. In fact, for very high orders n_2 or on a 32-bit computer the first scheme is better from the point of view of rounding errors. The reason is that in the second scheme for the evaluation of $P_{n_2}^{(1/2)}(\tilde{Q})$ we have to use the product form in terms of the roots ρ_j in (17). Even using the optimized ordering of roots defined in [1, 13], this is numerically less stable than the recursive evaluation according to (25), (31). If one uses the first scheme both $P_{n_2}^{(2)}$ in (12) and R_{n_3} in (14)-(15) can be evaluated recursively. Nevertheless, on a 64-bit machine both methods work well and in case of (19) the determination of the least-squares optimized polynomials is somewhat simpler.

The global accept-reject step for the gauge field can be performed after full sweeps over the gauge field links. A good choice for the order n_1 of the first polynomial $P_{n_1}^{(1)}$ is such that the average acceptance probability of the noisy correction be near 90%. One can decrease n_1 and/or increase the acceptance probability by updating only some subsets of the links before the accept-reject step. This might be useful on lattices larger than the largest lattice $12^3 \cdot 24$ considered in [5].

3.2 Measurement correction: reweighting

The multi-bosonic algorithms become exact only in the limit of infinitely high polynomial orders: $n \to \infty$ in (2) or, in the two-step approximation scheme, $n_2 \to \infty$ in (8). Instead of investigating the dependence on the polynomial order by performing several simulations, it is practically better to fix some high order for the simulation and perform another correction in the "measurement" of expectation values by still finer polynomials. This is done by *reweighting* the configurations in the measurement of different quantities. In case

6

of $N_f = 2$ flavours this kind of reweighting has been used in [10] within the polynomial hybrid Monte Carlo scheme. As remarked above, $N_f = 2$ is special because the reweighting can be performed by an iterative inversion. The general case can, however, also be treated by a further polynomial approximation.

The measurement correction for general N_f has been introduced in [14]. It is based on a polynomial approximation $P_{n_4}^{(4)}$ which satisfies

$$\lim_{n_4 \to \infty} P_{n_1}^{(1)}(x) P_{n_2}^{(2)}(x) P_{n_4}^{(4)}(x) = x^{-N_f/2} , \qquad x \in [\epsilon', \lambda] . \qquad (20)$$

The interval $[\epsilon', \lambda]$ can be chosen, for instance, such that $\epsilon' = 0, \lambda = \lambda_{max}$, where λ_{max} is an absolute upper bound of the eigenvalues of $Q^\dagger Q = \tilde{Q}^2$. In this case the limit $n_4 \to \infty$ is exact on an arbitrary gauge configuration. For the evaluation of $P_{n_4}^{(4)}$ one can use n_4-independent recursive relations (see section 4), which can be stopped by observing the convergence of the result. After reweighting the expectation value of a quantity A is given by

$$\langle A \rangle = \frac{\langle A \exp\{\eta^\dagger [1 - P_{n_4}^{(4)}(Q^\dagger Q)]\eta\}\rangle_{U,\eta}}{\langle \exp\{\eta^\dagger [1 - P_{n_4}^{(4)}(Q^\dagger Q)]\eta\}\rangle_{U,\eta}} , \qquad (21)$$

where η is a simple Gaussian noise like η' in (13). Here $\langle \ldots \rangle_{U,\eta}$ denotes an expectation value on the gauge field sequence, which is obtained in the two-step process described in the previous subsection, and on a sequence of independent η's. The expectation value with respect to the η-sequence can be considered as a Monte Carlo updating process with the trivial action $S_\eta \equiv \eta^\dagger \eta$. The length of the η-sequence on a fixed gauge configuration can be, in principle, arbitrarily chosen. In praxis it can be optimized for obtaining the smallest possible errors.

The application of the measurement correction is most important for quantities which are sensitive for small eigenvalues of the fermion matrix $Q^\dagger Q$. The polynomial approximations are worst near $x = 0$ where the function $x^{-N_f/2}$ diverges. In the exact effective gauge action, including the fermion determinant, the configuration with a small eigenvalue Λ are suppressed by $\Lambda^{N_f/2}$. The polynomials at finite order are not able to provide such a strong suppression, therefore in the updating sequence of the gauge fields there are more configurations with small eigenvalues than needed. The *exceptional configurations* with exceptionally small eigenvalues have to be supressed by the reweighting. This can be achieved by choosing $\epsilon' = 0$ and a high enough order n_4. It is also possible to take some non-zero ϵ' and determine the eigenvalues below it exactly. Each eigenvalue $\Lambda < \epsilon'$ is taken into account by an additional reweighting factor $\Lambda^{N_f/2} P_{n_1}^{(1)}(\Lambda) P_{n_2}^{(2)}(\Lambda)$. The stochastic correction in (21) is then restricted to the subspace orthogonal to these eigenvectors. Instead of $\epsilon' > 0$ one can also keep $\epsilon' = 0$ and project out a fixed number of smallest eigenvalues.

Let us note that, in principle, it would be enough to perform just a single kind of correction. But to omit the reweighting does not pay because it is much more comfortable to investigate the (small) effects of different n_4 values on the expectation values than to perform several simulations with increasing values of n_2. Without the updating correction the whole correction could be done by reweighting in the measurements. However, in practice this would not work either. The reason is that a first polynomial with relatively low order does not sufficiently suppress the exceptional configurations. As a consequence, the reweighting factors would become too small and would reduce the effective statistics considerably. In addition, the very small eigenvalues are changing slowly in the update and this would imply longer autocorrelations.

A moderate surplus of gauge configurations with small eigenvalues may, however, be advantageous because it allows for an easier tunneling among sectors with different topological charges. For small fermion masses on large physical volumes this is expected to be more important than the prize one has to pay for it by reweighting, provided that the reweighting has only a moderate effect.

4   Least-squares optimized polynomials

The basic ingredient of multi-bosonic fermion algorithms is the construction of the necessary optimized polynomial approximations. The least-squares optimization provides a general and flexible framework which is well suited for the requirements of multi-bosonic algorithms [13]. In the first part of this section the necessary basic formulae are collected. In the second part a simple example is considered: in case of an appropriately chosen weight function the least-squares optimized polynomials for the approximation of the function $x^{-\alpha}$ are expressed in terms of Jacobi polynomials.

4.1   Definition and basic relations

The general theory of least-squares optimized polynomial approximations can be inferred from the literature [16, 17]. Here we introduce the basic formulae in the way it has been done in [13] for the specific needs of multi-bosonic fermion algorithms. We shall keep the notations there, apart from a few changes which allow for more generality.

We want to approximate the real function $f(x)$ in the interval $x \in [\epsilon, \lambda]$ by a polynomial $P_n(x)$ of degree n. The aim is to minimize the deviation norm

$$\delta_n \equiv \left\{ N_{\epsilon,\lambda}^{-1} \int_{\epsilon}^{\lambda} dx \, w(x)^2 \left[f(x) - P_n(x) \right]^2 \right\}^{\frac{1}{2}} . \tag{22}$$

Here $w(x)$ is an arbitrary real weight function and the overall normalization factor $N_{\epsilon,\lambda}$ can be chosen by convenience, for instance, as

$$N_{\epsilon,\lambda} \equiv \int_\epsilon^\lambda dx\, w(x)^2 f(x)^2 \ . \tag{23}$$

A typical example of functions to be approximated is $f(x) = x^{-\alpha}/\bar{P}(x)$ with $\alpha > 0$ and some polynomial $\bar{P}(x)$. The interval is usually such that $0 \leq \epsilon < \lambda$. For optimizing the relative deviation one takes a weight function $w(x) = f(x)^{-1}$.

It turns out useful to introduce orthogonal polynomials $\Phi_\mu(x)$ ($\mu = 0, 1, 2, \ldots$) satisfying

$$\int_\epsilon^\lambda dx\, w(x)^2 \Phi_\mu(x) \Phi_\nu(x) = \delta_{\mu\nu} q_\nu \ . \tag{24}$$

and expand the polynomial $P_n(x)$ in terms of them:

$$P_n(x) = \sum_{\nu=0}^n d_{n\nu} \Phi_\nu(x) \ . \tag{25}$$

Besides the normalization factor q_ν let us also introduce, for later purposes, the integrals p_ν and s_ν by

$$q_\nu \equiv \int_\epsilon^\lambda dx\, w(x)^2 \Phi_\nu(x)^2 \ , \quad p_\nu \equiv \int_\epsilon^\lambda dx\, w(x)^2 \Phi_\nu(x)^2 x \ , \quad s_\nu \equiv \int_\epsilon^\lambda dx\, w(x)^2 x^\nu \ . \tag{26}$$

It can be easily shown that the expansion coefficients $d_{n\nu}$ minimizing δ_n are independent from n and are given by

$$d_{n\nu} \equiv d_\nu = \frac{b_\nu}{q_\nu} \ , \tag{27}$$

where

$$b_\nu \equiv \int_\epsilon^\lambda dx\, w(x)^2 f(x) \Phi_\nu(x) \ . \tag{28}$$

The minimal value of δ_n^2 is

$$\delta_n^2 = 1 - N_{\epsilon,\lambda}^{-1} \sum_{\nu=0}^n d_\nu b_\nu \ . \tag{29}$$

The above orthogonal polynomials satisfy three-term recurrence relations which are very useful for numerical evaluation. The first two of them with $\mu = 0, 1$ are given by

$$\Phi_0(x) = 1 \ , \qquad \Phi_1(x) = x - \frac{s_1}{s_0} \ . \tag{30}$$

The higher order polynomials $\Phi_\mu(x)$ for $\mu = 2, 3, \ldots$ can be obtained from the recurrence relation

$$\Phi_{\mu+1}(x) = (x + \beta_\mu)\Phi_\mu(x) + \gamma_{\mu-1}\Phi_{\mu-1}(x) \ , \qquad (\mu = 1, 2, \ldots) \ , \tag{31}$$

where the recurrence coefficients are given by

$$\beta_\mu = -\frac{p_\mu}{q_\mu} \ , \qquad \gamma_{\mu-1} = -\frac{q_\mu}{q_{\mu-1}} \ . \tag{32}$$

Using these relations on can set up a recursive scheme for the computation of the orthogonal polynomials in terms of the basic integrals s_ν defined in (26). Defining the polynomial coefficients $f_{\mu\nu}$ $(0 \leq \nu \leq \mu)$ by

$$\Phi_\mu(x) = \sum_{\nu=0}^{\mu} f_{\mu\nu} x^{\mu-\nu} \tag{33}$$

the above recurrence relations imply the normalization convention

$$f_{\mu 0} = 1 \;, \qquad (\mu = 0, 1, 2, \ldots) \;, \tag{34}$$

and one can easily show that q_μ and p_μ satisfy

$$q_\mu = \sum_{\nu=0}^{\mu} f_{\mu\nu} s_{2\mu-\nu} \;, \qquad p_\mu = \sum_{\nu=0}^{\mu} f_{\mu\nu} \left(s_{2\mu+1-\nu} + f_{\mu 1} s_{2\mu-\nu}\right) \;. \tag{35}$$

The coefficients themselves can be calculated from $f_{11} = -s_1/s_0$ and (31) which gives

$$
\begin{aligned}
f_{\mu+1,1} &= f_{\mu,1} + \beta_\mu \;, \\
f_{\mu+1,2} &= f_{\mu,2} + \beta_\mu f_{\mu,1} + \gamma_{\mu-1} \;, \\
f_{\mu+1,3} &= f_{\mu,3} + \beta_\mu f_{\mu,2} + \gamma_{\mu-1} f_{\mu-1,1} \;, \\
&\ldots \\
f_{\mu+1,\mu} &= f_{\mu,\mu} + \beta_\mu f_{\mu,\mu-1} + \gamma_{\mu-1} f_{\mu-1,\mu-2} \;, \\
f_{\mu+1,\mu+1} &= \beta_\mu f_{\mu,\mu} + \gamma_{\mu-1} f_{\mu-1,\mu-1} \;.
\end{aligned} \tag{36}
$$

The polynomial and recurrence coefficients are recursively determined by (34)-(36). The expansion coefficients for the optimized polynomial $P_n(x)$ can be obtained from (27) and

$$b_\mu = \sum_{\nu=0}^{\mu} f_{\mu\nu} \int_\epsilon^\lambda dx \, w(x)^2 f(x) x^{\mu-\nu} \;. \tag{37}$$

4.2   A simple example: Jacobi polynomials

The approximation interval $[\epsilon, \lambda]$ can be transformed to some standard interval, say, $[-1, 1]$ by the linear mapping

$$\xi = \frac{2x - \lambda - \epsilon}{\lambda - \epsilon} \;, \qquad x = \frac{\xi}{2}(\lambda - \epsilon) + \frac{1}{2}(\lambda + \epsilon) \;. \tag{38}$$

A weight factor $(1+\xi)^\rho (1-\xi)^\sigma$ with $\rho, \sigma > -1$ corresponds in the original interval to the weight factor

$$w^{(\rho,\sigma)}(x)^2 = (x - \epsilon)^\rho (\lambda - x)^\sigma \;. \tag{39}$$

Taking, for instance, $\rho = 2\alpha$, $\sigma = 0$ this weight is similar to the one for relative deviation from the function $f(x) = x^{-\alpha}$, which would be just $x^{2\alpha}$. In fact, for $\epsilon = 0$ these are exactly

the same and for small ϵ the difference is negligible. The advantage of considering the weight factor in (39) is that the corresponding orthogonal polynomials are simply related to the Jacobi polynomials [18, 19], namely

$$\Phi_\nu^{(\rho,\sigma)}(x) = (\lambda - \epsilon)^\nu \nu! \frac{\Gamma(\rho + \sigma + \nu + 1)}{\Gamma(\rho + \sigma + 2\nu + 1)} P_\nu^{(\sigma,\rho)} \left(\frac{2x - \lambda - \epsilon}{\lambda - \epsilon} \right) . \tag{40}$$

Our normalization convention (34) implies that

$$q_\nu^{(\rho,\sigma)} = (\lambda - \epsilon)^{\rho+\sigma+2\nu+1} \nu! \frac{\Gamma(\rho + \nu + 1)\Gamma(\sigma + \nu + 1)\Gamma(\rho + \sigma + \nu + 1)}{\Gamma(\rho + \sigma + 2\nu + 1)\Gamma(\rho + \sigma + 2\nu + 2)} . \tag{41}$$

The coefficients of the orthogonal polynomials are now given by

$$f_{\mu\nu}^{(\rho,\sigma)} = \sum_{\omega=0}^{\nu} (-\epsilon)^{\nu-\omega} (\epsilon - \lambda)^\omega \binom{\mu - \omega}{\nu - \omega} \binom{\mu}{\omega} \frac{\Gamma(\rho + \mu + 1)\Gamma(\rho + \sigma + 2\mu - \omega + 1)}{\Gamma(\rho + \mu - \omega + 1)\Gamma(\rho + \sigma + 2\mu + 1)} . \tag{42}$$

In particular, we have

$$f_{\mu 0}^{(\rho,\sigma)} = 1 , \qquad f_{11}^{(\rho,\sigma)} = -\epsilon - (\lambda - \epsilon)\frac{(\rho + 1)}{(\rho + \sigma + 2)} . \tag{43}$$

The coefficients β, γ in the recurrence relation (31) can be derived from the known recurrence relations of the Jacobi polynomials:

$$\beta_\mu^{(\rho,\sigma)} = -\frac{1}{2}(\lambda + \epsilon) + \frac{(\sigma^2 - \rho^2)(\lambda - \epsilon)}{2(\rho + \sigma + 2\mu)(\rho + \sigma + 2\mu + 2)} ,$$

$$\gamma_{\mu-1}^{(\rho,\sigma)} = -(\lambda - \epsilon)^2 \frac{\mu(\rho + \mu)(\sigma + \mu)(\rho + \sigma + \mu)}{(\rho + \sigma + 2\mu - 1)(\rho + \sigma + 2\mu)^2(\rho + \sigma + 2\mu + 1)} . \tag{44}$$

In order to obtain the expansion coefficients of the least-squares optimized polynomials one has to perform the integrals in (37). As an example, let us consider the function $f(x) = x^{-\alpha}$ when the necessery integrals can be expressed by hypergeometric functions:

$$\int_\epsilon^\lambda dx \, (x - \epsilon)^\rho (\lambda - x)^\sigma x^{\mu-\nu-\alpha} =$$

$$= (\lambda - \epsilon)^{\rho+\sigma+1} \lambda^{\mu-\nu-\alpha} \frac{\Gamma(\rho + 1)\Gamma(\sigma + 1)}{\Gamma(\rho + \sigma + 2)} F\left(\alpha - \mu + \nu, \sigma + 1; \rho + \sigma + 2; 1 - \frac{\epsilon}{\lambda}\right) . \tag{45}$$

Let us now consider, for simplicity, only the case $\epsilon = 0$, when we obtain

$$b_\mu^{(\rho,\sigma)} = (-1)^\mu \lambda^{1+\rho+\sigma+\mu-\alpha} \frac{\Gamma(\rho + \sigma + \mu + 1)\Gamma(\alpha + \mu)\Gamma(\rho - \alpha + 1)\Gamma(\sigma + \mu + 1)}{\Gamma(\rho + \sigma + 2\mu + 1)\Gamma(\alpha)\Gamma(\rho + \sigma - \alpha + \mu + 2)} . \tag{46}$$

Combined with (27) and (41) this leads to

$$d_\mu^{(\rho,\sigma)} = (-1)^\mu \lambda^{-\mu-\alpha} \frac{\Gamma(\rho + \sigma + 2\mu + 2)\Gamma(\alpha + \mu)\Gamma(\rho - \alpha + 1)}{\mu! \, \Gamma(\rho + \mu + 1)\Gamma(\alpha)\Gamma(\rho + \sigma - \alpha + \mu + 2)} . \tag{47}$$

11

These formulae can be used, for instance, for fractional inversion. For the parameters ρ, σ the natural choice in this case is $\rho = 2\alpha, \sigma = 0$ which corresponds to the optimization of the relative deviation from the function $f(x) = x^{-\alpha}$. As we have seen in section 4.1, the optimized polynomials are the truncated expansions of $x^{-\alpha}$ in terms of the Jacobi polynomials $P^{(2\alpha,0)}$. The Gegenbauer polynomials proposed in [20] for fractional inversion correspond to a different choice, namely $\rho = \sigma = \alpha - \frac{1}{2}$. This is because of the relation

$$C_n^\alpha(x) = \frac{\Gamma(n+2\alpha)\Gamma(\alpha+\frac{1}{2})}{\Gamma(2\alpha)\Gamma(n+\alpha+\frac{1}{2})} P_n^{(\alpha-\frac{1}{2},\alpha-\frac{1}{2})}(x) \ . \tag{48}$$

Note that for the simple case $\alpha = 1$ we have here the Chebyshev polynomials of second kind: $C_n^1(x) = U_n(x)$.

The special case $\alpha = \frac{1}{2}$ is interesting for the numerical evaluation of the zero mass lattice action proposed by Neuberger [21]. In this case, in order to obtain the least-squares optimized relative deviation with weight function $w(x) = x$, the function $x^{-\frac{1}{2}}$ has to be expanded in the Jacobi polynomials $P^{(1,0)}$. Note that this is different both from the Chebyshev and the Legendre expansions applied in [22]. The former would correspond to take $P^{(-\frac{1}{2},-\frac{1}{2})}$, the latter to $P^{(0,0)}$. The corresponding weight functions would be $[x(\lambda-x)]^{-\frac{1}{2}}$ and 1, respectively. As a consequence of the divergence of the weight factor at $x = 0$, the Chebyshev expansion is not appropriate for an approximation in an interval with $\epsilon = 0$. This can be immediately seen from the divergence of $d_\mu^{(-\frac{1}{2},-\frac{1}{2})}$ at $\alpha = \frac{1}{2}$ in (47).

The advantage of the Jacobi polynomials appearing in these examples is that they are analytically known. The more general least-squares optimized polynomials defined in the previous subsection can also be numerically expanded in terms of them. This is sometimes more comfortable than the entirely numerical approach.

5 Outlook

The two-step multi-bosonic algorithm has been shown to work properly in a recent large scale numerical simulation of light dynamical gluinos [4, 5]. Since gluinos are Majorana fermions the effective number of (Dirac-) fermion flavours is in this case $N_f = \frac{1}{2}$. This simulation is demanding because it deals with nearly zero and negative fermion masses in reasonably large physical volumes. It turned out possible to investigate the expected first order phase transition at zero gluino mass. The algorithm was able to cope with the metastability of phases near the phase transition. Up to now the investigated physical volumes were not very large: in case of the largest ($12^3 \cdot 24$) lattice the product of the spatial extension times the square root of the string tension was $L\sqrt{\sigma} \simeq 2.4$. The experience shows, however, that simulations with light gluinos and $L\sqrt{\sigma} \simeq 5$ or larger would be

feasible with reasonable effort. Another important conclusion in [5] is that the inclusion of the sign of the fermion determinant, actually Pfaffian for Majorana fermions, can be achieved by determining the spectral flow of the hermitian fermion matrix $\tilde{Q}$ as a function of increasing hopping parameter.

An interesting application of the two-step multi-bosonic algorithm is the numerical simulation of QCD. Up to now no serious effort was taken in this direction but first tests will be performed soon. A particularly relevant aspect for QCD is the possibility to deal with an odd number of quark flavours (N_f). The popular hybrid Monte Carlo algorithm can only be applied for even N_f. For non-even N_f one can use finite step-size molucular dynamics algorithms like the one in [23], but then the extrapolation to zero step size is an additional difficulty.

References

[1] I. Montvay, *Nucl. Phys.* B **466,** 259 (1996).

[2] M. Lüscher, *Nucl. Phys.* B **418,** 637 (1994).

[3] A.D. Kennedy, J. Kuti, *Phys. Rev. Lett.* **54,** 2473 (1985);
A.D. Kennedy, J. Kuti, S. Meyer, B.J. Pendleton, *Phys. Rev.* D **38,** 627 (1988).

[4] R. Kirchner, S. Luckmann, I. Montvay, K. Spanderen, J. Westphalen, *Phys. Lett.* B **446,** 209 (1999).

[5] I. Campos, A. Feo, R. Kirchner, S. Luckmann, I. Montvay, G. Münster, K. Spanderen, J. Westphalen, hep-lat/9903014.

[6] C. Alexandrou, A. Borrelli, P. de Forcrand, A. Galli, F. Jegerlehner, *Nucl. Phys.* B **456,** 296 (1995).

[7] A. Boriçi, P. de Forcrand, *Nucl. Phys.* B **454,** 645 (1995).

[8] A. Borrelli, P. de Forcrand and A. Galli, *Nucl. Phys.* B **477,** 809 (1996).

[9] C. Alexandrou, A. Borici, A. Feo, P. de Forcrand, A. Galli, F. Jegerlehner, T. Takaishi, *Nucl. Phys. Proc. Suppl.* B **53,** 435 (1997); *Nucl. Phys. Proc. Suppl.* B **63,** 406 (1998); hep-lat/9811028.

[10] R. Frezzotti, K. Jansen, *Phys. Lett.* B **402,** 328 (1997); hep-lat/9808011; hep-lat/9808038.

[11] U. Wolff, *Nucl. Phys. Proc. Suppl.* B **63,** 937 (1998).

[12] P. de Forcrand, hep-lat/9809145.

[13] I. Montvay, *Comput. Phys. Commun.* **109,** 144 (1998).

[14] R. Kirchner, S. Luckmann, I. Montvay, K. Spanderen, J. Westphalen, to appear in the proceedings of the Lattice '98 Conference, Boulder, July 1998, hep-lat/9808024.

[15] K. Spanderen, *Monte Carlo simulations of SU(2) Yang-Mills theory with dynamical gluinos,* PhD Thesis, University Münster, August 1998 (in German).

[16] L. Fox, I.B. Parker, *Chebyshev Polynomials in Numerical Analysis,* Oxford University Press, 1968.

[17] T.J. Rivlin, *An Introduction to the Approximation of Functions,* Blaisdell Publ. Company, 1969.

[18] G. Freud, *Orthogonale Polynome,* Birkhäuser Verlag, 1969.

[19] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products,* Academic Press, 1965.

[20] B. Bunk, *Nucl. Phys. Proc. Suppl.* B **63,** 952 (1998).

[21] H. Neuberger, *Phys. Lett.* B **417** 141 (1998).

[22] P. Hernandez, K. Jansen, M. Lüscher, hep-lat/9808010.

[23] S. Gottlieb, W. Liu, D. Toussaint, R.L. Renken, R.L. Sugar, *Phys. Rev.* D **35,** 2531 (1987).